AA214 – Solution to Problem Set Split/Factorization

1. *Multisplit Schemes*

   (a) In delta form the implicit Euler time differencing scheme is

   $$[I - hA_1 - hA_2 - hA_3 - hA_4]\Delta u_n = h[A_1 + A_2 + A_3 + A_4]u_n + hf + O(h^2)$$

   In factored delta form the implicit scheme becomes

   $$[I - hA_1]\,[I - hA_2]\,[I - hA_3]\,[I - hA_4]\Delta u_n = h[A_1 + A_2 + A_3 + A_4]u_n + hf + O(h^2)$$

   the error term that results from the factoring is

   $$\begin{aligned} er_f = {} & h^2[A_1A_2 + A_1A_3 + A_1A_4 + A_2A_3 + A_2A_4 + A_3A_4]\Delta u_n \\ & -h^3[A_1A_2A_3 + A_1A_2A_4 + A_1A_3A_4 + A_2A_3A_4]\Delta u_n \\ & +h^4A_1A_2A_3A_4\Delta u_n \end{aligned}$$

   This is $O(h^3)$, so the error term for the factored and unfactored forms is the same order, namely $O(h^2)$.

   (b) In delta form the explicit-implicit scheme is

   $$[I - hA_1 - hA_3]\Delta u_n = h[A_1 + A_2 + A_3 + A_4]u_n + hf + O(h^2)$$

   In factored delta form the explicit-implicit scheme becomes

   $$[I - hA_1]\,[I - hA_3]\Delta u_n = h[A_1 + A_2 + A_3 + A_4]u_n + hf + O(h^2)$$

   the error term that results from the factoring is

   $$er_f = h^2A_1A_3\Delta u_n$$

   This is $O(h^3)$, so again the error term for the factored and unfactored forms is the same order, namely $O(h^2)$.

   (c) In scalar form the fully implicit unfactored scheme is

   $$(1 - h\lambda_1 - h\lambda_2 - h\lambda_3 - h\lambda_4)(E - 1)u_n = h(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4)u_n + ha + O(h^2)$$

   Where E is the familiar shift operator. The characteristic polynomial, $P(E)$, and the particular polynomial, $Q(E)$, are, respectively

   $$\begin{aligned} P(E) &= (1 - h\lambda_1 - h\lambda_2 - h\lambda_3 - h\lambda_4)E - 1 \\ Q(E) &= h \end{aligned}$$

   Solving for the root of the characteristic polynomial gives

   $$\sigma = \frac{1}{1 - h\lambda_1 - h\lambda_2 - h\lambda_3 - h\lambda_4}$$

   The resulting stability, convergence, and accuracy are as follows:
   Stability. Assuming $\Re(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4) \le 0$ then $|\sigma| \le 1$ for all h; therefore the method is unconditionally stable.
   Convergence. Since

   $$\lim_{h\to\infty} |\sigma| = 0$$

   the method is rapidly convergent for large values of h.

Accuracy. Assuming convergence, the steady state solution is

$$u_s = a\frac{Q(1)}{P(1)} = -\frac{a}{\lambda_1+\lambda_2+\lambda_3+\lambda_4}$$

This is the correct steady state solution.
In scalar form the fully implicit factored scheme is

$$(1-h\lambda_1)(1-h\lambda_2)(1-h\lambda_3)(1-h\lambda_4)(E-1)u_n = h(\lambda_1+\lambda_2+\lambda_3+\lambda_4)u_n + ha + O(h^2)$$

The characteristic polynomial, $P(E)$, and the particular polynomial, $Q(E)$, are, respectively

$$\begin{aligned} P(E) &= (1-h\lambda_1)(1-h\lambda_2)(1-h\lambda_3)(1-h\lambda_4)(E-1) - h(\lambda_1+\lambda_2+\lambda_3+\lambda_4) \\ Q(E) &= h \end{aligned}$$

Solving for the root of the characteristic polynomial gives

$$\sigma = \frac{(1-h\lambda_1)(1-h\lambda_2)(1-h\lambda_3)(1-h\lambda_4) + h(\lambda_1+\lambda_2+\lambda_3+\lambda_4)}{(1-h\lambda_1)(1-h\lambda_2)(1-h\lambda_3)(1-h\lambda_4)}$$

The resulting stability, convergence, and accuracy are as follows:
Stability. Assuming $\Re(\lambda_1+\lambda_2+\lambda_3+\lambda_4) \le 0$ then $|\sigma| \le 1$ for some h. For some h and λ's the method can be catastrophically unstable. Suppose, for example, that $\lambda_1 = 1$ and that $\lambda_1+\lambda_2+\lambda_3+\lambda_4 \neq 0$, then, if $h = 1$, $|\sigma|$ will be infinite.
Convergence. Since

$$\lim_{h\to\infty} |\sigma| = 1$$

care must be taken in selecting a time step size which will ensure that $|\sigma| < 1$. As long as $\Re(\lambda_1+\lambda_2+\lambda_3+\lambda_4) < 0$ there will always exist an h for which $|\sigma| < 1$, however, for some sets of λ's even the optimum h will produce $|\sigma|$ only slightly less than 1, in which case the convergence rate will be abysmal.
Accuracy. Assuming convergence, the steady state solution is

$$u_s = a\frac{Q(1)}{P(1)} = -\frac{a}{\lambda_1+\lambda_2+\lambda_3+\lambda_4}$$

This is the correct steady state solution.

(d) In scalar form the explicit-implicit unfactored scheme is

$$(1-h\lambda_1-h\lambda_3)(E-1)u_n = h(\lambda_1+\lambda_2+\lambda_3+\lambda_4)u_n + ha + O(h^2)$$

The characteristic polynomial, $P(E)$, and the particular polynomial, $Q(E)$, are, respectively

$$\begin{aligned} P(E) &= (1-h\lambda_1-h\lambda_3)E - 1 - h\lambda_2 - h\lambda_4 \\ Q(E) &= h \end{aligned}$$

Solving for the root of the characteristic polynomial gives

$$\sigma = \frac{1+h\lambda_2+h\lambda_4}{1-h\lambda_1-h\lambda_3}$$

The resulting stability, convergence, and accuracy are as follows:
Stability. Assuming $\Re(\lambda_1+\lambda_2+\lambda_3+\lambda_4) \le 0$ then $|\sigma| \le 1$ for some h. Scenarios can be constructed in which the method will be catastrophically unstable.

Convergence. Since

$$\lim_{h\to\infty} |\sigma| = -\frac{\lambda_2 + \lambda_4}{\lambda_1 + \lambda_3}$$

the method may or may not be rapidly convergent. If $\lambda_2 + \lambda_4 = 0$ and $\lambda_1 + \lambda_3 \neq 0$, then the method will be rapidly convergent for large values of h; otherwise, there will exist a noninfinite optimum h for which $|\sigma|$ will be less than one, but, perhaps, only slightly.

Accuracy. Assuming convergence, the steady state solution is

$$u_s = a\frac{Q(1)}{P(1)} = -\frac{a}{\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4}$$

This is the correct steady state solution.

In scalar form the explicit-implicit factored scheme is

$$(1 - h\lambda_1)(1 - h\lambda_3)(E - 1)u_n = h(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4)u_n + ha + O(h^2)$$

The characteristic polynomial, $P(E)$, and the particular polynomial, $Q(E)$, are, respectively

$$\begin{aligned} P(E) &= (1 - h\lambda_1)(1 - h\lambda_3)E - 1 - h\lambda_2 - h\lambda_4 - h^2\lambda_1\lambda_3 \\ Q(E) &= h \end{aligned}$$

Solving for the root of the characteristic polynomial gives

$$\sigma = \frac{1 + h\lambda_2 + h\lambda_4 + h^2\lambda_1\lambda_3}{(1 - h\lambda_1)(1 - h\lambda_3)}$$

The resulting stability, convergence, and accuracy are as follows:

Stability. Assuming $\Re(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4) \leq 0$ then $|\sigma| \leq 1$ for some h. Scenarios can be constructed in which the method will be catastrophically unstable.

Convergence. Since

$$\lim_{h\to\infty} |\sigma| = 1$$

care must be taken in selecting a time step size which will ensure that $|\sigma| < 1$. As long as $\Re(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4) < 0$ there will always exist an h for which $|\sigma| < 1$, however, for some sets of λ's even the optimum h will produce $|\sigma|$ only slightly less than 1, in which case the convergence rate will be abysmal.

Accuracy. Assuming convergence, the steady state solution is

$$u_s = a\frac{Q(1)}{P(1)} = -\frac{a}{\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4}$$

This is the correct steady state solution.

2. *System Splitting, Plus-Minus Splitting*

   (a) The coupled set of PDE can be expressed as

   $$\frac{\partial}{\partial t}\begin{bmatrix} u \\ v \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}\frac{\partial}{\partial x}\begin{bmatrix} u \\ v \end{bmatrix} = 0$$

   defining the vector q and the matrix A as follows

   $$q = \begin{bmatrix} u \\ v \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$$

   the coupled set of PDE can be expressed in matrix-vector form

   $$\frac{\partial q}{\partial t} + A\frac{\partial q}{\partial x} = 0$$

(b) From the characteristic equation for A the eigenvalues are obtained

$$\lambda^2 - 1 = 0 \longrightarrow \lambda_1 = -1,\ \lambda_2 = 1$$

The eigenvectors can be readily obtained

$$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{21} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \longrightarrow \begin{bmatrix} x_{11} \\ x_{21} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} x_{12} \\ x_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \longrightarrow \begin{bmatrix} x_{12} \\ x_{22} \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

The eigenvalue, eigenvector, and inverse eigenvector matrices are then

$$\Lambda = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \quad X = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \quad \text{and} \quad X^{-1} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

Applying a plus-minus split to the eigenvalue matrix yields

$$\Lambda^+ = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad \Lambda^- = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}$$

A plus-minus splitting of the A matrix can now be defined

$$A^+ \equiv X\Lambda^+X^{-1} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \quad \text{and} \quad A^- \equiv X\Lambda^-X^{-1} = \frac{1}{2} \begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix}$$

A plus-minus splitting of the flux vector E ($\equiv Aq$) can also be defined

$$E^+ \equiv A^+q = \frac{1}{2} \begin{bmatrix} u - v \\ v - u \end{bmatrix} \quad \text{and} \quad E^- \equiv A^-q = -\frac{1}{2} \begin{bmatrix} u + v \\ u + v \end{bmatrix}$$

(c) Applying implicit Euler time differencing to the system yields

$$\frac{1}{h}(q_{n+1} - q_n) + A\delta_x q_{n+1} + O(h) = 0$$

Multiplying through by h and rearranging terms gives

$$[I + hA\delta_x]q_{n+1} = q_n + O(h^2)$$

Applying plus-minus splitting to the system, and noting that $\delta_x E^+_{n+1}$ ($= A^+\delta_x q_{n+1}$) is stable to backward differencing, and that $\delta_x E^-_{n+1}$ ($= A^-\delta_x q_{n+1}$) is stable to forward differencing, gives

$$\left[I + hA^+\delta_x^b + hA^-\delta_x^f\right] q_{n+1} = q_n + O(h^2)$$

In delta form this is

$$\left[I + hA^+\delta_x^b + hA^-\delta_x^f\right] \Delta q_n = -\left[hA^+\delta_x^b + hA^-\delta_x^f\right] q_n + O(h^2)$$

In factored delta form this is

$$\left[I + hA^+\delta_x^b\right]\left[I + hA^-\delta_x^f\right] \Delta q_n = -\left[hA^+\delta_x^b + hA^-\delta_x^f\right] q_n + O(h^2)$$

the error term that results from the factoring is

$$er_f = h^2 A^+ A^- \delta_x^b \delta_x^f \Delta q_n$$

This is $O(h^3)$, so, timewise, the scheme is still first order accurate. After spatially discretizing, and assuming first order spatial derivatives are used, the matrix operator for the unfactored form will be block tridiagonal (the blocks will be 2×2's); for the factored form it will be, in a block sense, a lower triangular matrix times an upper triangular matrix (or vice versa), with each matrix having a block bandwidth of two. In this spatially one-dimensional case the computational savings of factorization won't be great since the multiplication and division operations for solving the two forms for $m \times m$ size matrices will be of the same order in m. For two or three spatial dimensions cases, however, factorization can yield multiplication and division operation counts that are lower order in m.